COMPUTATIONAL STUDY OF ANTIBIOTIC RESISTANCE IN DIARRHEAL PATHOGENS AND IDENTIFICATION OF POTENTIAL DRUG TARGETS AND VACCINE CANDIDATES

SYNOPSIS

Submitted in fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

IN

BIOINFORMATICS

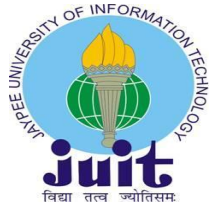
BY

KUSUM

Enrollment Number- 136503

Under the Supervision of

DR. JAYASHREE RAMANA



DEPARTMENT OF BIOTECHNOLOGY & BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT, SOLAN-173234, HP, INDIA

Introduction

Diarrhea is a manifestation of intestinal dysfunction that results in increased stool frequency, characterized by loss of water, electrolytes, and/or nutrients. Diarrheal diseases rank second only to pneumonia as the leading cause of deaths among children under five years of age worldwide [1, 2]. Despite the fact that mortality rate of diarrheal diseases was substantially reduced by almost 31.1% from 1·8 million in 2000 to 1·3 million in 2013 [3] due to inception of many immunization programs for diarrheal disease control, the proportional mortality rate still is soaring with an estimated 7,60,000 under 5 children deaths and 1.7 billion diarrheal incidences annually [4]. In India, diarrheal diseases account for approximately 13% of all childhood deaths under age 5; additionally contributing to another 1% of the neonatal deaths [5]. Diarrheal pediatric death toll is higher than that of AIDS, measles, and malaria combined. Diarrhea has been classified into three categories based on clinical presentations [6].

1.) Acute Watery Diarrhea (Non-Inflammatory) which lasts for several hours or days.

2.) Acute Bloody Diarrhea termed as dysentery (Inflammatory).

3.) Persistent Diarrhea (Enteric Fever) which lasts for 14 days or longer.

Diarrhea is caused by a wide spectrum of infectious organisms including viruses, bacteria, protozoa, and helminths. These diarrheal agents follow diverse pathogenesis mechanisms and vary in their route of transmission, inoculum size requirement and host preferences [7]. Several medication categories are available to deal with different aspects of diarrheal diseases which include opiates, antibiotics, bile acid sequestrants and heavy metals. Table 1 lists currently available antibiotics and vaccines against diarrheal pathogens. Zinc supplementation as an adjunct to ORS is recommended for diarrheal management in malnourished children and in persistent diarrhea particularly in developing countries [8].

Antibacterial resistance (ABR)

ABR which is the ability of the bacteria to develop resistance to the antibiotics to which it was originally susceptible has emerged as a global health concern threatening the treatment of ever-widening range of infectious diseases. From 1940-1980, average resistance time for a drug was exceeding ten years, nevertheless in present scenario it has reduced to one year [9]. Efforts invested in research and development for antibiotics discovery have reduced significantly in last few decades.

Deaths attributable to drug resistance would be concentrated more in South-East Asia region by 2050 which is as high as 4,730,000 [10]. India regarded as the largest consumer of antibiotics in the world was reported with a whooping 43% increase in antibiotic use from 2000 to 2010 [11]; where 18% of the drugs were used without prescription [12]. In context of diarrheal pathogens, intermediate to high level of resistance has been reported globally.

Table 1: List of currently available drugs and vaccines against diarrheal pathogens.

Organism	Drugs	Vaccines		
		Live attenuated	Killed whole-cell vaccines	Subunit
Shigella	Ciprofloxacin, Pivmecillinam, Azithromycin	CVD1208S, WRSS1	ETEC whole cell	Chemical glycoconjugate, Recombinant glycoconjugate, Synthetic glycoconjugate, InvaplexAR, GMMA, OMV, DB Fusion, PSSP-1, 34 kDa OMP
Salmonella	Azithromycin, Fluoroquinolones, Ceftriaxone	JOL919, JOL1364 (Chickens)		
Campylobacter	Erythromycin, Azithromycin and Fluoroquinolones			
Enterotoxigenic *E. coli* (ETEC)	Ciprofloxacin and Azithromycin	ACE527, ZH9	ETVAX	Dukoral®(CTB), dmLT, ST, LT patch, Fimbrial tip adhesion, CF/CS consensus peptide fused to dmLT-ST mutants, YghJ, EatA
Rotavirus		RotaTeq®, Rotarix, Rotavac		
V. cholerae	Doxycycline and Tetracycline	CVD 103-HgR, Mutacol	Dukoral, Shanchol	
E. histolytica	Metronidazole, Tinidazole, Nitroimidazole			
G. lamblia	Metronidazole and Ornidazole	α1- Giardin		

Mutational resistance plays a critical role in continuous bacterial evolution and has clinical significance [13]. SNPs proved to be potential biomarkers due to their genome-wide availability in virtually all organisms and being large in numbers. SNPs find applications in genetic mapping to decipher breeding pedigree, in evolutionary studies to identify genomic divergence aimed at explicating the role of speciation in evolution and in genome wide association studies to correlate genotypic variations with the phenotypic traits [14]. The identification and cataloguing of SNPs implicated in antibiotic resistance and resistance genes in a public repository proves to be a useful resource for various purposes in both basic and applied research. Large scale studies have been

undertaken globally to provide a catalog of antibiotic resistance associated mutations from different pathogens such as MTCID, MGDD, CRAD, ARDB etc [15-18]. However, there have been no efforts to curate the experimental SNPs reported to be associated with antibiotic resistance in diarrheal pathogens which is scattered all over the web. We have designed DBDiaSNP database which a comprehensive repository of resistance genes and antibiotic resistance associated mutations curated from published literature. The efficient utilization of SNP based web resources involves understanding how individual genetic variations are decoded into structural modifications in protein drug targets which get translated into differential therapeutic drug responses. We have performed molecular modeling and molecular dynamics simulation studies to understand the effect of mutations in gyrase A subunit on quinolone affinity in ETEC and *C. jejuni.* With ever-widening range of microbes being resistant to existing antibiotics, bioterrorism threat and emergence of new infectious diseases, there remains a dire need to reinvigorate the antibiotic pipeline. We have utilized the power of omics approaches to identify potential drug targets and vaccine candidates against various diarrheal pathogens.

Objectives:

- **To develop a database of mutations and resistance genes from various diarrheal pathogens**
- **To study the structural signatures of the pathogen mutations deposited in the database and their consequences on drug binding landscape**
- ***In silico* prediction of potential therapeutic drug targets and vaccine candidates against a wide variety of diarrheal pathogens**

Objective 1: To develop a database of mutations and resistance genes from various diarrheal pathogens

Methods

Database System Implementation

DBDiaSNP has been designed on MySQL 5.1 that provides multi-layered server design with fast execution times. The IIS server handles queries from web clients through PHP scripts to perform searches which allow researchers to obtain SNP information. An overview of the DBDiaSNP architecture is given in Figure 1.a. As shown in Figure 1.b, the database is a comprehensive resource of drug-resistance mutations and resistance genes in diarrheal pathogens and hosts. We conducted a systematic review to identify antibiotic resistance mutations and resistance genes from literature to be included in the database. The variations are categorized into SNPs, substitutions, frameshift mutations, deletions, insertions. For each mutation, the database provides complete codon change at nucleotide or amino acid level. Resistance genes reported in different pathogens can also be browsed and are categorized into Plasmid mediated, Transposons mediated, Integrons and gene Cassettes mediated resistance modes according to its subcellular location. Table 2, Table 3 and Table 4 show the database statistics which houses 590 mutations and 326 resistance genes associated with diarrheal pathogens (viruses, protozoans and bacterial species) and 356 mutations associated with hosts. DBDiaSNP allows users to search for SNP and resistance gene information based on organism of interest which includes both host and pathogens. The search results provide related SNP and resistance gene data and their annotations. The information on corresponding protein, PDB identifiers, function etc. has been extracted from different web resources dbSNP, OMIM, ClinVar, PDB [19-22]; thus providing a comprehensive compilation of the SNPs and resistance genes reported in the clinical isolates of various diarrheal pathogens.

Results and Discussion

By providing a comprehensive repository of drug resistance mutation and resistance genes, we hope it will open new avenues to the discoveries that will have applications from diagnostics to drug discovery. There is pressing need to develop tools that can identify the antibiotic-resistant pathogens rapidly and with high accuracy. SNPs based diagnostic methods detect mutations associated with a particular disease, but such tools extensively rely on information about the mutations and their relative frequency [23].

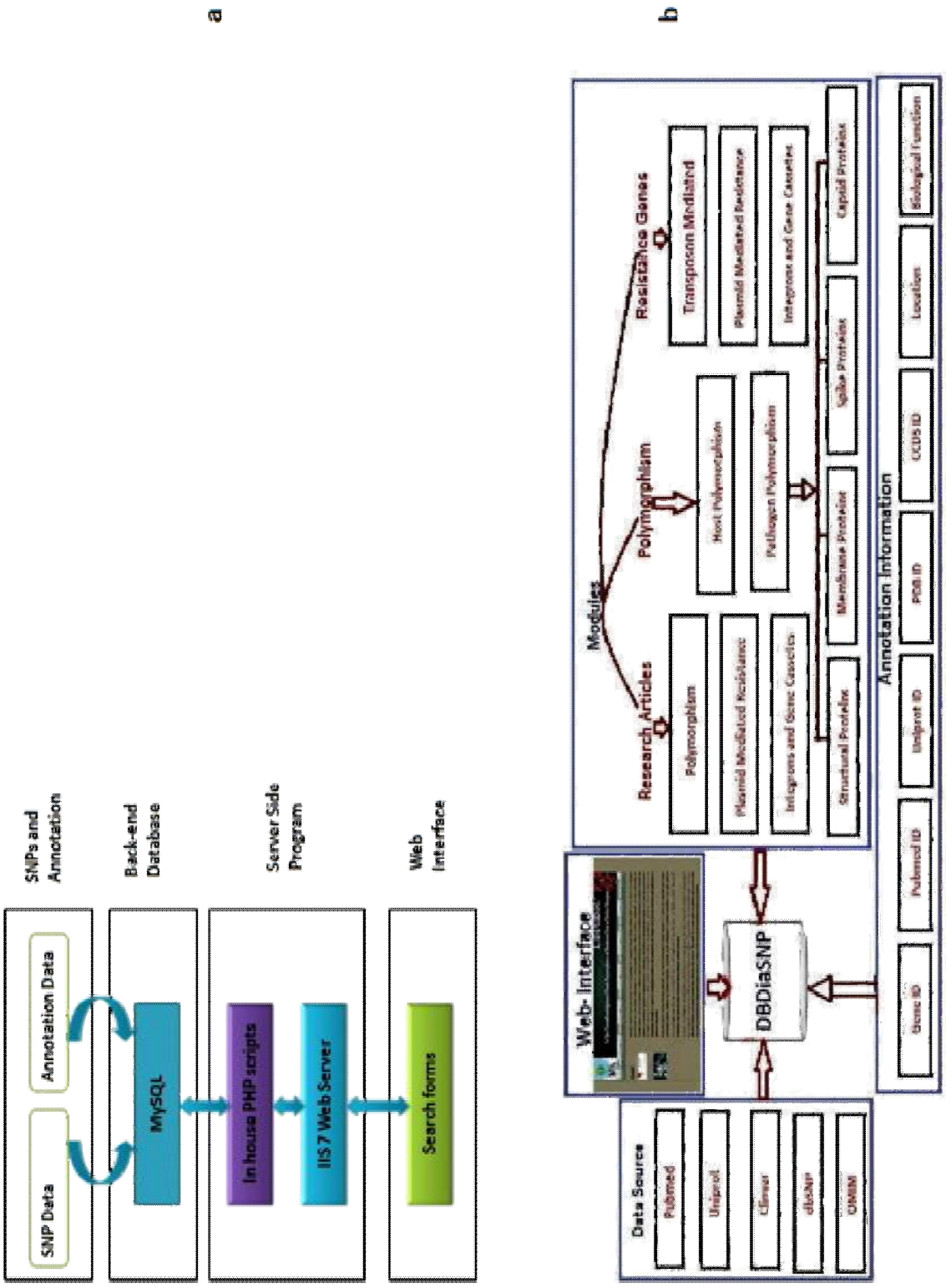


Figure 1: DBDiaSNP architecture overview. **a)** Standard three-tier architecture of DBDiaSNP; **b)** Organizational structure of the DBDiaSNP.

Table 2: SNPs statistics along with genes: Species wise distribution in pathogens

Organism	#SNPs (Genes)
Viruses	221 (10)
Escherichia coli	91 (15)
Entamoeba histolytica	11 (1)
Campylobacter spp.	98 (4)
Shigella spp.	13 (3)
Vibrio cholerae	16 (1)
Aeromonas caviae	4 (2)
Clostridium difficile	46 (4)
Helicobacter pylori	71 (3)
Salmonella spp.	19 (3)
Total	**590 (46)**

Table 3: Resistance Gene statistics: Species wise distribution in pathogens

Organism	#Resistance Genes
Escherichia coli	68
Enterococcus spp.	10
Campylobacter spp.	10
Shigella spp.	58
Vibrio spp.	60
Aeromonas spp.	26
Clostridium spp.	10
Klebsiella spp.	16
Salmonella spp.	68
Total	326

Table 4: SNPs statistics along with genes: Species wise distribution in hosts

Organism	#SNPs (Genes)
Humans	291 (57)
Porcine	65 (3)
Total	**356 (60)**

Mutations arising as a result of treatment with different drug analogs may help explicate the geographic differences in drug effectiveness [24]. The database will aid in lead discovery against antibiotic resistant pathogens by mapping the mutations onto the protein structure. We have created an interactive web interface to retrieve information about mutations and resistance genes in diarrheal pathogens associated with drug resistance.

Conclusions

DBDiaSNP is intended to pace up with research efforts for global surveillance and control of diarrhea. The variations reported in the database represent potential markers of antibiotic resistance in diarrheal pathogens which might help in understanding genotype-phenotype relationships. For future translational research involving integrative biology and global health, the presented database offers veritable potentials, particularly for developing countries and worldwide monitoring and personalized effective treatment of pathogens associated with diarrhea.

Objective 2: To study the structural signatures of the pathogen mutations deposited in the database and their consequences on drug binding landscape

Methodology

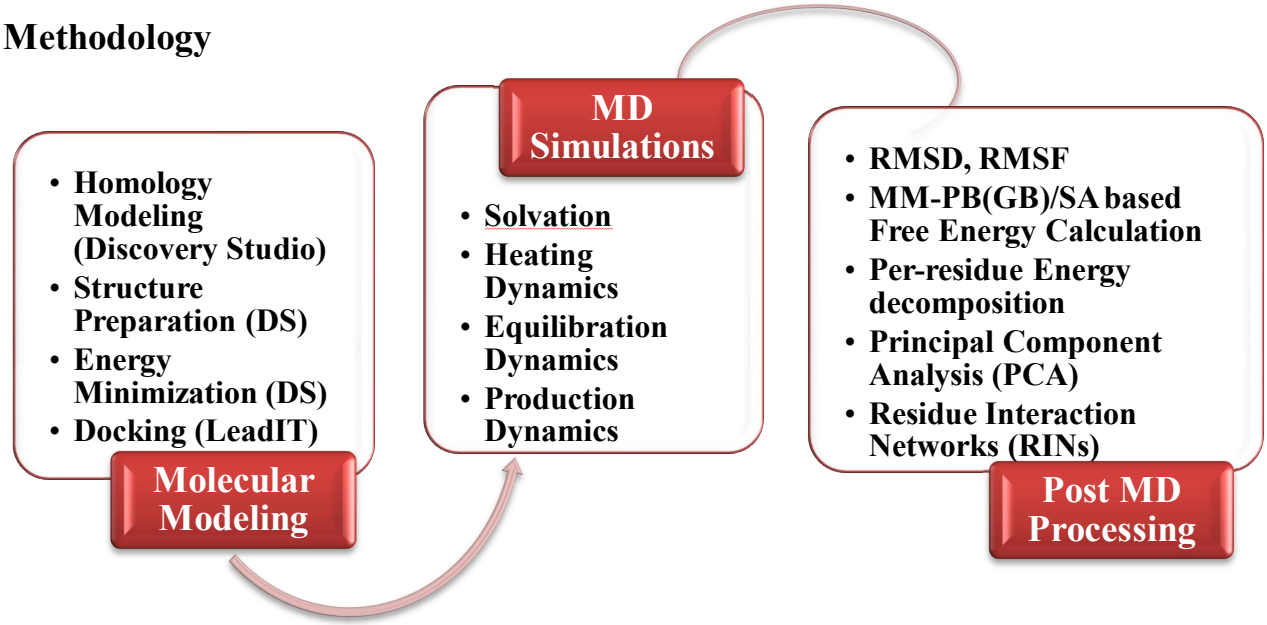


Figure 3: Molecular Modeling and MD simulations protocol for studying effect of mutations on ciprofloxacin binding in gyrA in *C. jejuni* and Enterotoxigenic *E. Coli*.

Results and Discussion

Molecular Docking

Ciprofloxacin interacting with WT gyrA through hydrogen bonds with residues Lys42, Asp87, and Arg91 predicted a docking score of -25.50. In MT Ser83Leu and double mutant Ser83Leu/Asp87Asn docking score was observed to be -21.2327 and -18.8967 involving hydrogen bonds with Lys42, Asp87, and Arg91 (Figure 4). In WT gyrA Arg91 bonded with ciprofloxacin atoms O2 and O4 while in the MT gyrA hydrogen bond with ciprofloxacin O4 was lost (Table 5), thereby contributing to a decrease in the docking score. Asp87 being directly hydrogen bonded to the ciprofloxacin H43 atom, plays a crucial role in the quinolone binding with gyrA. In the double mutant hydrogen bond length for residue 87 was increased from 2.76 Å in WT gyrA to 3.25 Å in significantly decreasing the bond energy.

Molecular Dynamics Simulation

To evaluate the conformational flexibility of the WT and MT structures, time dependent root-mean-square deviations (RMSD) of the backbone atoms from the corresponding energy minimized

structure were monitored. Figure 5 depicts RMSDs of WT and MT models compared to the reference structure as a function of time. RMSDs were calculated from the last 5 ns of the 10-ns MD simulations. The trajectories for all the simulation systems were stabilized over the time course, with moderately small fluctuations in the models. It was proposed that these mutations do not disrupt the backbone conformation or cause significant global change. In order to assess the effect of these mutation on dynamic behavior of residues, root mean square fluctuations (RMSFs) between native and S83L, and S83L/D87N were generated and plotted in Figure 6. The deviations in residue fluctuations may influence the interaction with ligand and thus the ligand binding efficiency. The deviations in residue fluctuations may influence the interaction with ligand and thus the ligand binding efficiency. The regions of high RMSFs were not confined to vicinity of the mutated residues only, but were distributed across entire protein backbone. The difference in RMSDs of WT and MT models at several apex points was higher than 1 Å which further strengthen our argument that such mutations introduced in native structure influence the dynamic properties of the protein.

Flexibility of active site residue Tyr 122 is reduced in MT as compared to WT indicating the relative rigidity of this region in MT models. To evaluate the free energy differences between WT and MT, MM-PB(GB)/SA methods were employed. Predicted binding energies and energy component breakdown for different simulation systems are summarized in Table 6. The total binding free energies using both methods follows the same trend with a reduction in the binding affinity in MT relative to WT. Reduced binding affinities due to the mutations influence the efficiency of ciprofloxacin binding to gyrA.

Table 5: Docking scores and hydrogen bond residues involved in ciprofloxacin binding with WT and MT gyrA

Mutations	Score	Receptor Residue	Receptor Atom	Ligand Atom	Bond Lengths (Å)
Wild Type	-25.50	Lys 42	NZ	O3	3.04
		Arg 91	NH1	O4	2.79
		Arg 91	NH1	O2	3.01
		Asp 87	OD2	H43	2.76
Ser83Leu	-21.23	Lys 42	NZ	O3	2.74
		Arg 91	NH1	O2	2.75
		Asp 87	OD2	H43	2.80
Ser83Leu/Asp87Asn	-18.90	Lys 42	NZ	O3	2.96
		Lys 42	NZ	O4	2.73
		Arg 91	NH1	O2	2.75
		Asn 87	OD1	H42	3.25

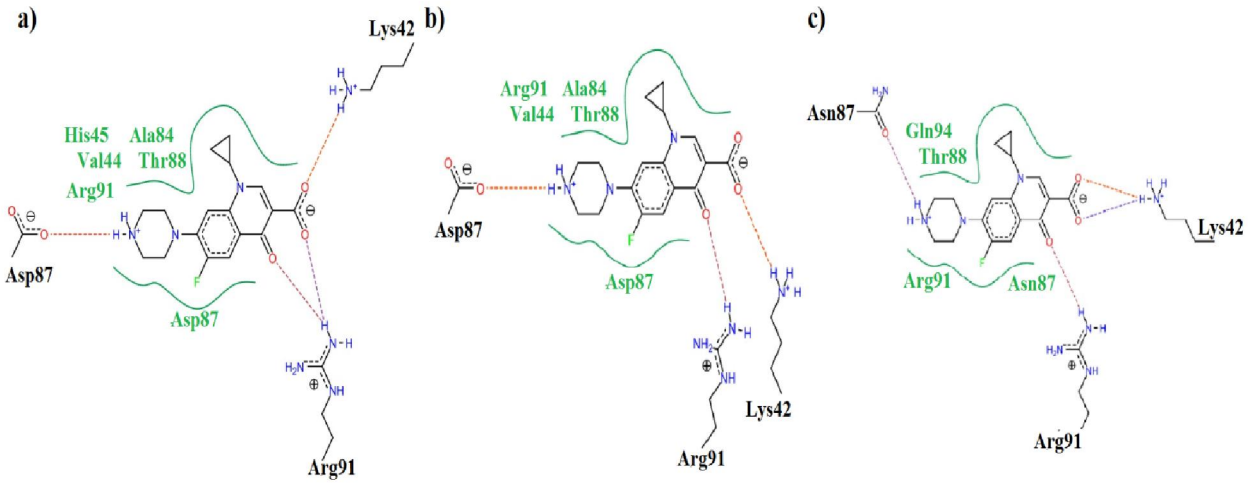


Figure 4: Hydrogen bonds involved in binding of ciprofloxacin with **a)** wild type **b)** mutated S83L and **c)** S83L/D87N QRDR of gyrA.

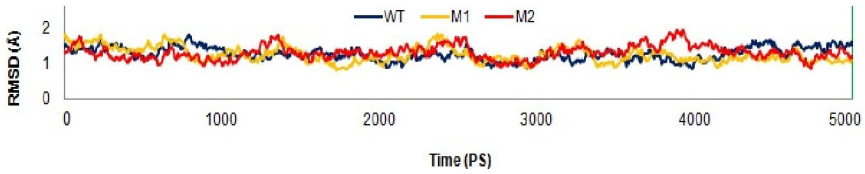


Figure 5: Time evolution RMSD of gyrA backbone throughout the 10 ns MD simulation time for WT (Blue), S83L mutant (Yellow), S83L/D87N mutant (Red).

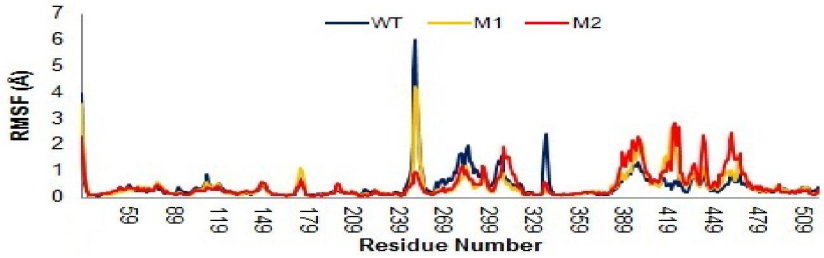


Figure 6: Time evolution RMSF of gyrA backbone atoms throughout the 10 ns MD simulation time for WT (Blue), S83L mutant (Yellow), S83L/D87N mutant (Red).

The predicted van der walls (ΔE_{VDW}) and electrostatic contribution (ΔE_{ELE}) are significantly higher in WT gyrA than both MT (Table 6). Per-residue decomposition of binding energy sheds light on the residues which majorly contributes to gyrA-ciprofloxacin binding in both WT and MT models. Figure 7 compares the protein–ligand interaction spectra between the WT and MT models and shows contribution of individual hot-spot residues to polar and non-polar interaction terms. Decrease in the overall interaction energy in MT models is largely accountable to change in the van der Waals contribution of hot-spot residues.

Table 6: MM-PB(GB)/SA based free binding energy profile of ciprofloxacin complexed with the wild and the S83L, S83L/D87N mutant types of gyrA.

Contribution	WT	S83L	S83L/D87N
ΔE_{INT}	0	0	0
ΔE_{VDW}	-24.2421	-13.4977	-5.6295
ΔE_{ELE}	-110.813	-40.6163	-33.5151
$\Delta E_{GAS/\ \Delta E\ MM}$	-135.055	-54.114	-39.1446
ΔG_{SOL-NP}	-2.5072	-0.8743	0.1548
ΔG_{PB}	123.253	45.0921	34.9382
$\Delta G_{SOLV,PB}$	120.746	44.2177	35.0931
$\Delta G_{ELE,PB}$	12.44	4.4758	1.4231
$H_{TOT,PB}$	-14.3093	-9.8963	-4.0515
ΔG_{GB}	121.811	46.6201	36.6038
$\Delta G_{SOLV,GB}$	118.514	44.7953	35.8509
$\Delta G_{ELE\ GB}$	10.9982	6.0038	3.0887
$H_{TOT,\ GB}$	-16.5413	-9.3187	-3.2937

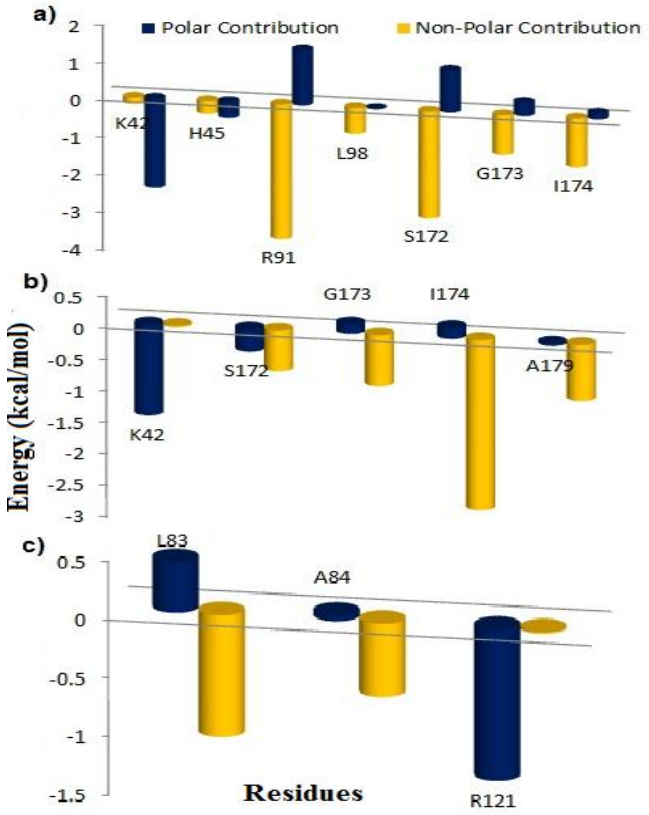


Figure 7: Polar and non-polar contribution of hot-spot residues from MM-PBSA based energy decomposition analysis of total binding energy (kcal/mol) during last 5 ns of MD simulation.

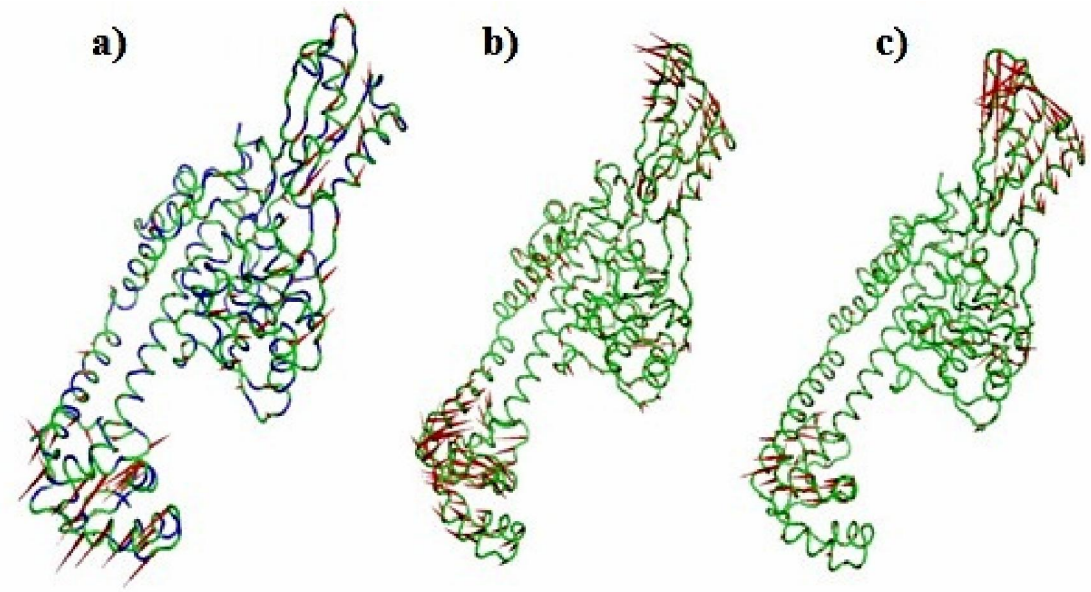


Figure 8: Porcupine plot of the first eigenvector obtained by PCA from simulation ensembles of (a) WT, (b) S83L mutant, and S83L/D87N mutant.

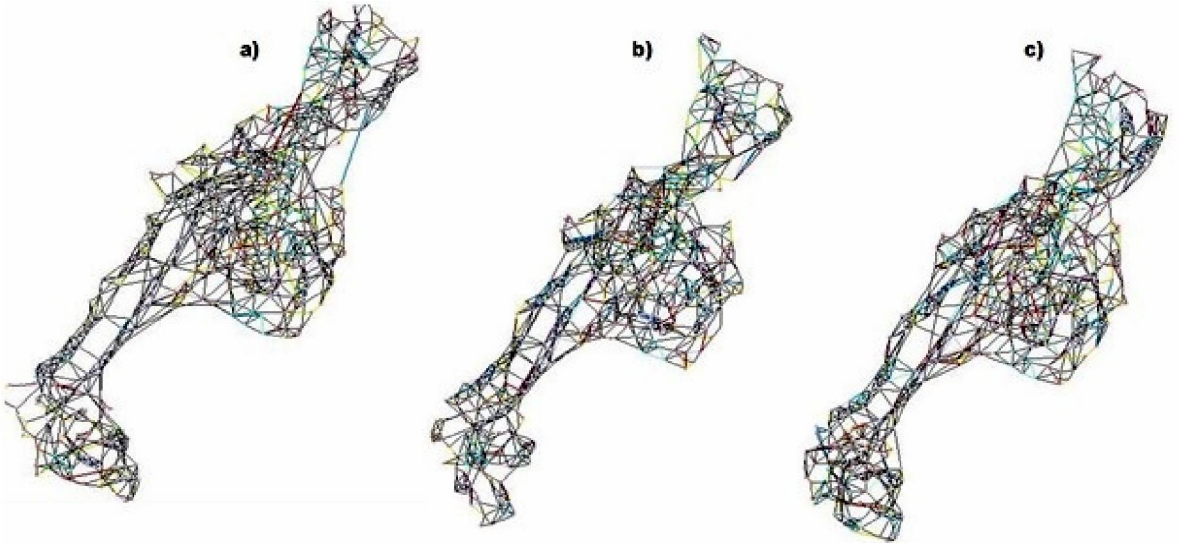


Figure 9: Residue interaction networks (RINs) comparison between a) WT, b) S83L mutant, and c) S83L/D87N mutant.

Nodes represented as: Helix Loop Sheet Default and

Edges represented as: Hydrogen Bond Peptide Bond Salt Bridge

Ionic bond Pi Stacking Pi-Cation Closest Atoms Interaction

Figure 8 shows porcupine plot for the ciprofloxacin bound WT and MT complexes along the direction of first principal component which shows significant difference in the overall pattern of global motions between three systems. Dominant motions represented by high magnitude arrows were mostly observed in the three distinctive domains of N-terminal regions (NTD). Consequently, overall conformations of trajectories from the last 5 ns simulation were different across different structures. To identify key residue interactions and explore the differences between different WT and MT models, RINs were generated using the representative structures from the last 5 ns trajectories of each molecular system. As evident from the RINs plot (Figure 9), RIN of MT variants of gyrA vary considerably when compared to the WT and each other. Though effect of changes at individual residue site may not be prominent, but collectively these changes in residue interactions could have extreme effect on protein structure altering the active site conformation.

Conclusion

Rapidly evolving computational platforms have made investigating the structural consequences of the mutations easier. Though WT and MT trajectories were exploring the same conformational space, however mutations significantly impact the dynamic behavior of complexes which was reflected by major disruptions in the residue interaction network and altered dominant motions which could destabilize active-site conformation reducing ligand binding efficiency. This study offers a comprehensive picture of genotype–phenotype association of quinolone resistance associated SNPs in gyrA.

Objective 3: *In silico* prediction of potential therapeutic drug targets and vaccine candidates against a wide variety of diarrheal pathogens

Enterotoxigenic *E. coli* (ETEC)

Methods

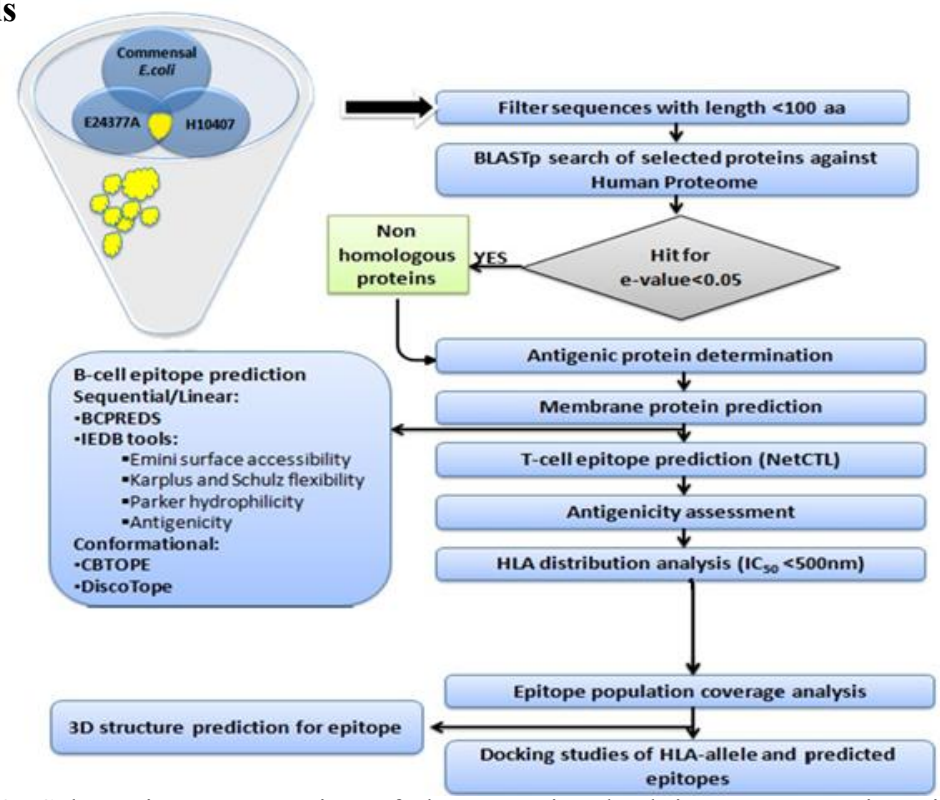


Figure 10: Schematic representation of the steps involved in target protein prioritization and antigenic epitope mapping.

Results and Discussion

The design of an epitope-based vaccine candidate against Enterotoxigenic E. coli has already been proposed but these studies are focused on some specific proteins like autotransporter protein, adhesion proteins like eae (intimin), colonization factors like CFA/1, CS2 and CS3 [25, 26]. However the genome of ETEC strains may encode several unveiled antigenic proteins, which have so far not been explored as vaccine candidates. In this study, we have used integrated comparative genomics and immunoinformatics analysis of available ETEC genomes in the search for vaccine candidates. The proteins were selected based on physicochemical properties like surface localization (secreted), or having transmembrane (TM) domains. The VaxiJen server predicted the protective antigens based on the overall antigenicity score and assisted to prioritize the proteins. For T-cell

epitope predictions, parameters such as T-cell processivity, the number of human leukocyte antigens (HLA) alleles covered, and significant population coverage were considered. The characteristics features of B-cell epitopes include physicochemical propensities like flexibility, hydrophilicity, solvent accessibility. First, the shared protein pool between E24377A and H10407 ETEC strains was determined by a reciprocal BLAST analysis. Next, proteins present in the commensal *E. coli* strain SE11 genomes were eliminated from this shared pool. Protein sequences from this shared pool were filtered based on length criterion. We subjected these protein sequences to a homology search against Human proteome database using BLASTP program from standalone blast suite and retrieved pathogen proteins which were non-homologous to humans. Identifying proteins non-homologous to humans is essential as it excludes the possibility of the peptide vaccine targeting hosts enzymes, thus avoiding adverse effects on humans [27]. This analysis rendered 61 protein sequences.

Table 7: Predicted potential T cell epitopes, along with their interacting MHC-I and MHC-II alleles with an affinity of <500 nM and corresponding IC50 values (bracketed)

Epitope	# MHC-peptide binders	MHC-I alleles	MHC-II alleles
LICFFTLSY	24	HLA-A*01:01(92.47), HLA-A*02:06 (105.68), HLA-A*02:17 (130.82), HLA-A*03:01 (305.49), HLA-A*11:01 (257.22), HLA-A*29:02 (30.89), HLA-A*30:02 (400.37), HLA-A*32:01 (468.75), HLA-A*32:07 (23.72), HLA-A*32:15 (164.53), HLA-A*68:01 (59.57), HLA-A*68:02 (492.04), HLA-A*68:23 (99.25), HLA-A*80:01 (26.42), HLA-B*15:01 (54.83), HLA-B*15:02 (293.24), HLA-B*15:03 (56.51), HLA-B*15:17 (212.18), HLA-B*27:20 (14.40), HLA-B*35:01 (133.53), HLA-B*40:13 (179.03), HLA-C*12:03 (14.73), HLA-C*14:02 (468.40)	HLA-DRB1*01:01 (9.40)
PLNPLILLY	23	HLA-A*01:01 (285.76), HLA-A*02:03 (4.80), HLA-A*02:17 (280.32), HLA-A*03:01 (323.59), HLA-A*11:01 (144.88), HLA-A*29:02 (25.63), HLA-A*31:01 (210.38), HLA-A*32:07 (20.42), HLA-A*32:15 (47.78), HLA-A*68:01 (127.64), HLA-A*68:23 (91.57), HLA-A*80:01 (2.30), HLA-B*15:02 (144.62), HLA-B*27:20 (4.76), HLA-B*35:01 (436.52), HLA-B*40:13(399.88), HLA-C*05:01 (102.02), HLA-C*07:01 (326.72), HLA-C*07:02 (295.24), HLA-C*12:03 (48.11), HLA-C*14:02 (228.36)	HLA-DRB1*01:01 (1.01), HLA-DRB1*04:01 (359.75)
PIVNLFLLY	22	HLA-A*01:01 (14.13), HLA-A*02:01 (456.04), HLA-A*02:03 (280.54), HLA-A*02:06 (38.55), HLA-A*03:01 (20.94), HLA-A*11:01 (49.32), HLA-A*26:01 (450.41), HLA-A*26:02 (108.22), HLA-A*29:02 (28.96), HLA-A*32:07 (26.92), HLA-A*32:15 (52.63), HLA-A*68:01 (99.54), HLA-A*68:23 (76.34), HLA-A*80:01 (4.53), HLA-B*15:02 (159.31), HLA-B*27:20 (35.69), HLA-B*35:01 (251.77), HLA-B*40:13 (93.53), HLA-C*03:03 (255.63), HLA-C*12:03 (23.62)	HLA-DRB1*01:01 (0.78), HLA-DRB1*07:01 (47.64)
SVSVFIFLF	27	HLA-A*02:01 (306.90), HLA-A*02:02 (269.78), HLA-A*02:06 (307.61), HLA-A*02:50 (74.00), HLA-A*11:01 (6.90), HLA-A*24:02 (467.34), HLA-A*24:03 (388.63), HLA-A*26:02 (63.58), HLA-A*29:02 (325.66), HLA-A*32:01 (111.41), HLA-A*32:07 (7.91), HLA-A*32:15 (179.99), HLA-A*68:01 (50.23), HLA-A*68:02 (202.77), HLA-A*68:23 (11.72), HLA-B*15:02 (222.45), HLA-B*15:03 (495.55), HLA-B*15:17 (93.69), HLA-B*27:20 (70.39), HLA-B*35:01 (242.66), HLA-B*40:13 (106.40), HLA-B*58:01 (454.26), HLA-C*03:03 (92.17), HLA-C*07:01 (239.98), HLA-C*07:02 (297.97), HLA-C*12:03 (84.96)	HLA-DRB1*01:01 (12.45)
SVFIFLFIY	25	HLA-A*02:01 (316.96), HLA-A*02:02 (467.74), HLA-A*02:06 (77.62), HLA-A*02:17 (90.50), HLA-A*03:01 (239.33), HLA-A*11:01 (18.84), HLA-A*26:01 (454.58), HLA-A*26:02 (73.17), HLA-A*29:02 (21.27), HLA-A*30:02 (77.89), HLA-A*32:01 (253.48), HLA-A*32:07 (13.09), HLA-A*32:15 (53.12), HLA-A*68:01 (15.56), HLA-A*68:02 (283.79), HLA-A*68:23 (5.53), HLA-A*80:01 (18.37), HLA-B*15:02 (157.12), HLA-B*15:17 (46.96), HLA-B*27:20 (41.83), HLA-B*35:01 (160.91), HLA-B*40:13 (20.37), HLA-C*07:02 (444.81), HLA-C*12:03 (11.07)	HLA-DRB1*01:01 (41.59)

The VaxiJen server used to assess the antigenicity of the protein sequences predicted 9 proteins as antigenic above a threshold of 0.6 were further scrutinized for their cellular location. It revealed that

3 proteins were membrane localized by combined predictions of PSORTb and TMHMM which were further used for epitope mapping. NetCTL predicted T-cell epitopes from each sequence against MHC-supertypes. 50 epitopes were predicted from the membrane localized antigenic sequences. These epitopes were further assessed by VaxiJen for antigenicity. For each high priority epitope, SMM based IEDB MHC I processing prediction tool retrieved the MHC-I alleles with IC50 value less than 500 nM which were potential epitope binders (Table 7). For population coverage analysis, epitopes along with their associated MHC-I and MHC-II alleles were input to IEDB population coverage analysis tool. Immune response elicitation of the 97.58% and 95.57% total world population was covered by the two predicted epitopic sets.

Table 8: Five most potential B-cell epitopes by combined predictions of AAP, BCPred, IEDB tools (Emini Surface Accessibility, Karplus and Schulz flexibility, Parker hydrophilicity), further filtered based on their antigenicity values (VaxiJen Score)

Uniprot Id	Start Residue	End Residue	Epitope	Length	VaxiJen Score
A7ZGR5	345	352	PETHKSDN	8	1.1952
	237	246	FKNQFNKKIT	10	0.9161
	179	189	YSSTKNIHQQK	11	1.0160
A7ZTH5	272	279	YSHDNTRT	8	1.9700
	303	319	EKRAEKIHELEEKEPRL	17	1.0362

Table 9: Potential conformational B-cell epitope residues predicted using CBTOPE and DiscoTope server, along with the corresponding secondary structure conformation each residue adopts and DiscoTope scores.

Uniprot Id	CBTOPE					DiscoTope		
	Start Residue	End Residue	Epitope	Length	Secondary Structure Type	Residue No	Residue	DiscoTope Score
A7ZGR5	322	322	I	1	C	238	K	2.081
	342	345	VLFP	4	CCCC	325	Q	1.718
	349	351	KSD	3	CCC	327	I	1.920
	404	405	IR	2	EC	328	N	3.520
						363	L	2.118
						366	K	1.393
A7ZGK4	17	17	V	1	C			
	136	139	PDKR	4	CCCC			
A7ZTH5	291	293	GLK	3	HHH			
	324	324	P	1	C			
	410	410	S	1	C			

Using Hex, different conformations of the predicted epitopes were generated and the best conformation was selected based on binding affinity scores. As per the criteria set for prediction of B-cell epitopes, the sequential epitopes were predicted using combined predictions of BCPREDS

server and IEDB tools; further filtered based on antigenicity values from VaxiJen. Antigenic regions common to both BCPred and AAP algorithms at BCPREDS server were subjected to IEDB Emini surface accessibility tool to predict peptides which were surface exposed. Predicted peptides were checked for flexibility and hydrophilicity using IEDB tools Karplus & Schulz flexibility prediction, and Parker hydrophilicity prediction. Sequential B-cell epitopes with antigenicity score >0.9 were selected (Table 8). Antibody interacting residues in membrane localized proteins predicted using CBTOPE server are summarized in Table 12 along with their secondary structure conformation. The probability of finding Conformational B cell epitopes is more in the coiled or turn region [28]. Majority of the antibody interacting residues we identified were also adopting coil conformation. Models predicted from I-TASSER were used by DiscoTope server. Residues which are predicted to be a part of conformational B-cell epitopes by DiscoTope server (score>1) are listed in Table 9.

C. jejuni Drug Target Identification

Methods:

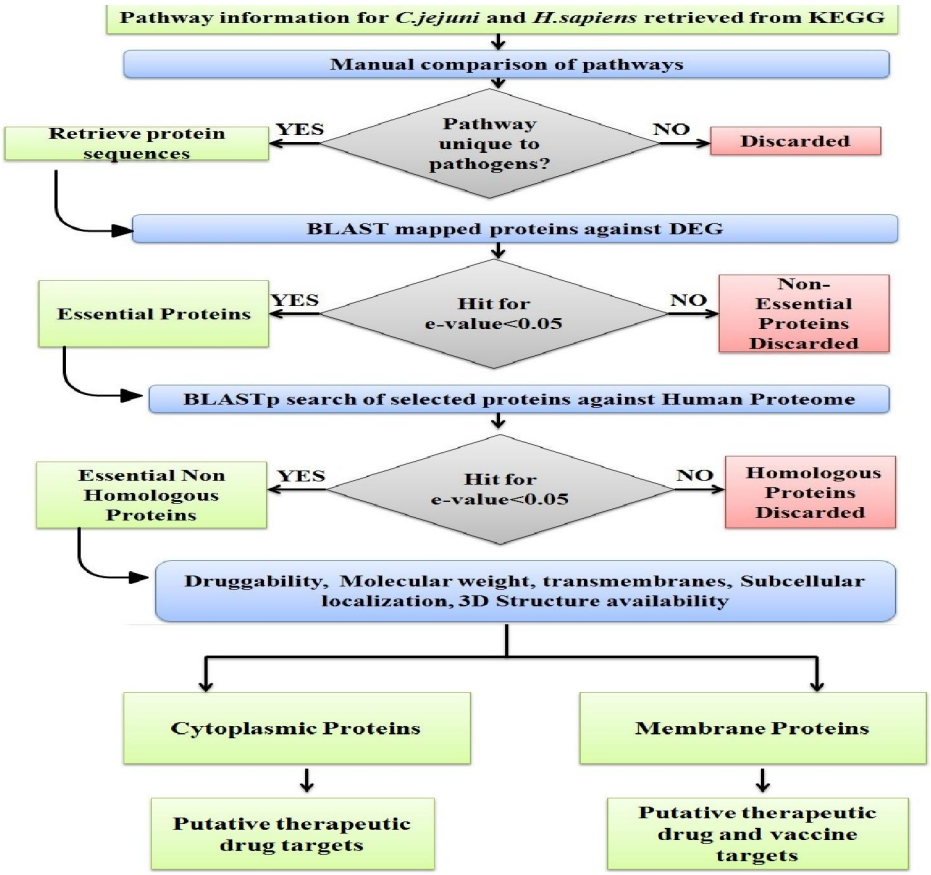


Figure 11: Schematic workflow of comparative genomics approach to identify drug targets in pathogenic *C. jejuni* strain NCTC11168

Results and Discussion

As described in the workflow (Figure 11) manual comparison of host and pathogen pathways was conducted which resulted in identification of 16 unique pathways to the pathogen. A total of 326 unique proteins were identified to be involved in 16 unique pathways. A BLASTP search of these 326 protein sequences against 551 essential genes from *C. jejuni* and *H. pylori* in DEG was performed which revealed a total of 115 essential protein sequences with hits below e-value 0.05. These essential protein sequences were further filtered out by homology search against human proteome for identification of non-homologous protein sequences. This comparison detected 66 essential non-homologous proteins with no hits against *H. sapiens* below e-value 0.05, a snippet of which is shown in Table 10.

Prioritization of Essential Non-homologous proteins for therapeutic drug targets

Although all 66 essential non-homologous identified proteins are potential drug targets, yet these can be further filtered using additional parameters which maximize the likelihood of landing to the best therapeutic target against pathogen. Subcellular localization prediction using PSORTb program identified 63 proteins to be of cytoplasmic origin and all the essential non-homologous proteins were found to be <110 kDa; suggesting possibility of experimental verification of the most of the identified targets. A search for availability of 3D structure identified only 6 proteins with experimentally solved structures which points to a gap in the structural characterization of pathogen proteins. Protein structural information can be used to a significant advantage in drug identification and validation [29].

Druggability of Potential Drug Targets

The proteins for which drugs are already available can prove to be useful starting points for drug discovery. 34 *C. jejuni* proteins were identified which shared high similarity to the binding partners of these drug targeted proteins from DrugBank. Finally, we have narrowed down the search of therapeutic targets to final 14 prioritized drug targets and vaccine candidates. pfs enzyme catalyzes the direct conversion of aminodeoxyfutalosine (AFL) into dehypoxanthine futalosine (DHFL) and adenine via the hydrolysis of the N-glycosidic bond which represents an essential step in menaquinone biosynthesis pathway [30]. Inhibition of alr and ddl leads to effective inhibition of peptidoglycan synthesis and ddl has been proposed to be an attractive drug target in *Mycobacterium tuberculosis* [31]. Both frdC and ccoP are important constituents of oxidative phosphorylation pathway for ATP formation. secD , tatC and Cj1094c transport large proteins across membranes.

Transport proteins are associated with pathogenesis and virulence [32, 33]. Peptidoglycan biosynthesis, the pathway with largest distribution of final identified drug targets can be exploited for therapeutic drug targets owing to its multiple target enzymes whose inhibition could lead to disruption of cell well and in turn attenuate bacterial cell growth. Majority of antibiotics in clinical use today target later steps of peptidoglycan synthesis [34]

Table 10 : Essential non-homologous proteins of *C. jejuni* with results of prioritization analysis (transmembrane and cellular location prediction)

No	Gene Name	Protein Id	Pathway Involved	Protein existence	Subcellular location	TMHMM	3D Exp	3D Model	MW (Da)
1	alr	Q9PP26	D-Alanine metabolism; Vancomycin resistance	Inferred	Cytoplasmic	No	No	Yes	37,256
2	ddl	Q9PPC2	D-Alanine metabolism; Vancomycin resistance; Peptidoglycan biosynthesis	Inferred	Cytoplasmic	No	No	Yes	39,908
3	lpxA	Q9PIM1	Lipopolysaccharide biosynthesis	Predicted	Cytoplasmic	No	Yes	Yes	28,651
4	lpxD	Q9PHC0	Lipopolysaccharide biosynthesis	Inferred	Unknown	No	No	Yes	34,679
5	lpxC	Q9PIZ5	Lipopolysaccharide biosynthesis	Inferred	Cytoplasmic	No	No	Yes	32,383
6	murA	Q9PP65	Peptidoglycan biosynthesis	Inferred	Cytoplasmic	No	No	Yes	45,186
7	murB	Q9PM01	Peptidoglycan biosynthesis	Inferred	Cytoplasmic	No	No	Yes	28,994
8	murC	Q9PNN7	Peptidoglycan biosynthesis	Inferred	Cytoplasmic	No	No	Yes	48,372
9	murF	Q0PA95	Peptidoglycan biosynthesis; Vancomycin resistance	Predicted	Cytoplasmic Membrane	Yes	No	Yes	55,234
10	pbpA	Q0PB07	Peptidoglycan biosynthesis; Beta-Lactam resistance	Predicted	Cytoplasmic Membrane	Yes	No	Yes	72,628
11	fba	Q0PAS0	Methane metabolism; Biosynthesis of secondary metabolites; Microbial metabolism in diverse environments	Predicted	Cytoplasmic	No	Yes	Yes	38,730
12	aroQ	Q9PJ53	Biosynthesis of secondary metabolites	Predicted	Cytoplasmic	No	No	Yes	17,595
13	pfs	Q0PC20	Biosynthesis of secondary metabolites	Predicted	Cytoplasmic	No	No	Yes	25,225
14	panD	Q9PIK3	Biosynthesis of secondary metabolites	Predicted	Cytoplasmic	No	Yes	Yes	13,974
15	panC	Q9PIK2	Biosynthesis of secondary metabolites	Predicted	Cytoplasmic	No	Yes	Yes	32,087
16	panB	Q9PIK1	Biosynthesis of secondary metabolites	Inferred	Cytoplasmic	No	No	Yes	30,202
17	pheA	Q0PBJ3	Biosynthesis of secondary metabolites	Predicted	Cytoplasmic	No	No	Yes	40,477
18	aroE	Q0PBA5	Biosynthesis of secondary metabolites	Inferred	Cytoplasmic	No	No	Yes	29,978
19	trpC	Q9PI11	Biosynthesis of secondary metabolites	Inferred	Cytoplasmic	No	No	Yes	29,573
20	ispH	P0C632	Biosynthesis of secondary metabolites	Inferred	Cytoplasmic	No	No	No	31,551
21	aroA	P0C630	Biosynthesis of secondary metabolites	Inferred	Cytoplasmic	No	No	Yes	47,288
22	aroB	Q9PNT2	Biosynthesis of secondary metabolites	Inferred	Cytoplasmic	No	No	Yes	39,596
23	dxr	Q9PMV3	Biosynthesis of secondary metabolites	Inferred	Cytoplasmic	No	No	Yes	39,548
24	Cj1370	Q0P8N9	Biosynthesis of secondary metabolites	Predicted	Cytoplasmic	No	No	Yes	17,035
25	aroC	Q9PM41	Biosynthesis of secondary metabolites	Predicted	Cytoplasmic	No	Yes	Yes	39,252
26	dapB	Q9PIT2	Biosynthesis of secondary metabolites; Microbial metabolism in diverse environments	Inferred	Cytoplasmic	No	No	Yes	25,696
27	rpiB	Q0P9X4	Biosynthesis of secondary metabolites; Microbial metabolism in diverse environments	Predicted	Cytoplasmic	No	No	Yes	16,293
28	pbpB	Q0PAZ1	Peptidoglycan biosynthesis; Beta-Lactam resistance	Predicted	Cytoplasmic Membrane	Yes	No	Yes	68,378
29	pbpC	Q0PAL6	Peptidoglycan biosynthesis; Beta-Lactam resistance	Predicted	Cytoplasmic Membrane	Yes	No	Yes	67,971
30	ftsW	Q0PBA2	Biosynthesis of secondary metabolites; Microbial metabolism in diverse environments; Two-component system	Predicted	Cytoplasmic Membrane	Yes	No	Yes	30,028

(continued)

S. typhimurium Drug Target Identification

Methods:

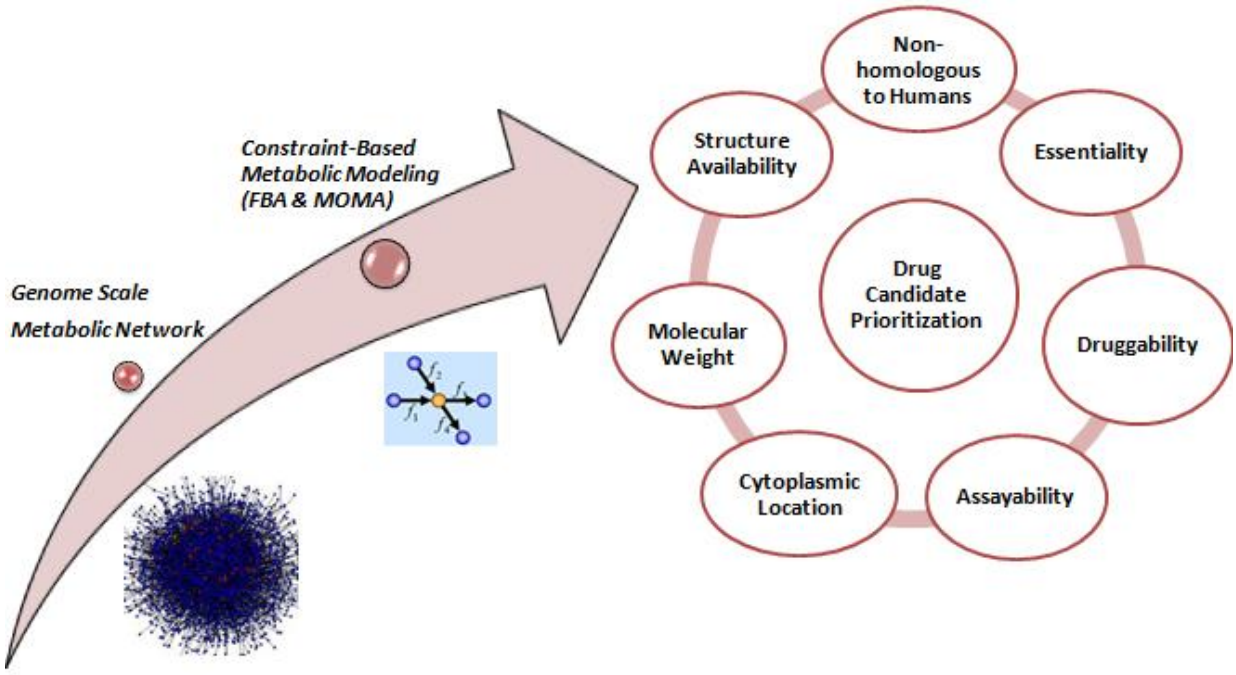


Figure 12: Schematic representation of the steps involved in target protein prioritization from genome scale metabolic network of *S. typhimurium* LT2.

Results and Discussion

For identifying the genes which are indispensable to the bacterial growth, we used standard FBA and MOMA models implemented in OptFlux as outlined in the workflow chart in Figure 12. We identified 73 genes to be critical using gene knockout studies. These genes rendered a reduction of greater than 90 percent in the total biomass yield during.

Identification of Essential Non-homologous Pathogen Proteins

A BLASTP search of the proteins corresponding to these critical genes by homology search against human proteome rendered 65 genes as non-homologous to *H. sapiens* below e-value 0.05. These proteins were further filtered by performing a BLASTP search against experimentally predicted essential genes of *Salmonella enterica* serovar *Typhi*, *Salmonella enterica* serovar *Typhi* Ty2, *Salmonella enterica* serovar *Typhimurium* SL1344, *Salmonella enterica* subsp. *enterica* serovar *Typhimurium* str. 14028S, and *Salmonella typhimurium* LT2 in DEG. This search revealed that a

total of 54 proteins were in alignment with experimentally verified essential genes from *Salmonella* spp.

Prioritization of drug targets

The essential gene set retrieved from this study was further prioritized through a multistep process which involved following criteria: a) molecular weight b) availability of experimental data on target crystal structure; c) subcellular localization; d) presence of transmembrane helices; e) druggability potential; and f) Assayability; to be reduced to a set of candidate drug targets. All the 54 proteins retrieved in previous step were having molecular weight <110 kDa which makes them amenable to isolation and experimental verification. In a search for 3D structural data, it was observed that crystal structure information was available for yadF, ribH, fabD, aroD, rfbG, rfbD, ppnK, and ribB proteins. Even after a decade of *S. typhimurium* LT2 genome being sequenced out, structural information is typically lacking for the proteins encoded by this strain. Typically majority of the drug targets were predicted to be of cytoplasmic origin as per PSORTb calculations. Subcellular localization of the drug target can shed some light on the protein function. With combined predictions of PSORTb and TMHHM server, 6 proteins *viz.* mraY, cdsA, pgsA, plsC, rfaL were predicted to be membrane localized with one or more helices traversing the membrane, opening up new avenues to explore these proteins as vaccine candidates. A snippet of these results is provided in Table 11. A BLASTp search of all the selected proteins was performed against the druggable proteins deposited in the DrugBank database to evaluate the druggability potential of the selected proteins. With combined predictions of PSORTb and TMHHM server, 6 proteins *viz.* mraY, cdsA, pgsA, plsC, rfaL were predicted to be membrane localized with one or more helices traversing the membrane, opening up new avenues to explore these proteins as vaccine candidates. A BLASTp search of all the selected proteins was performed against the druggable proteins deposited in the DrugBank database to evaluate the druggability potential of the selected proteins. We identified 47 proteins with a significantly higher similarity to the binders of these drug targeted proteins. For each drug target we have listed the identified drug types. Druggability potential is an important concern in drug target identification considering the fact that only 10-14% of the human genome is speculated to be druggable. Information on assayability of each protein was referenced from BRENDA database. Assayability information is critical to support the screening programs for establishing suitable lead compounds. After applying all the filtration criteria, we selected proteins involved in L-rhamnose biosynthesis, peptidoglycan biosynthesis, fatty acid biosynthesis, folate biosynthesis as our serving set of top priority therapeutic drug candidates.

Table11 :Similar binding partners of drugs available in Drug Bank inferred by homology search, druggabiltv, assayability, α-helices and pathway information on prioritized drug target proteins of *S. typhimurium* LT2

Gene	Drugg-ability	Drug Class	Essen-tiality	Location	TM-HMM	Assay-ability	Pathways Involved
ribF	✓	Experimental, Nutraceutical	✓	Cytoplasmic		✓	Riboflavin metabolism, Biosynthesis of secondary metabolites
folA	✓	Approved, Experimental, Nutraceutical	✓	Periplasmic		✓	Folate biosynthesis
murE	✓	Experimental	✓	Cytoplasmic		✓	Peptidoglycan biosynthesis, Lysine biosynthesis
murF	✓	Experimental	✓	Cytoplasmic		✓	Peptidoglycan biosynthesis, Vancomycin resistance, Lysine biosynthesis
mraY			✓	Cytoplasmic Membrane	10	✓	Peptidoglycan biosynthesis, Vancomycin resistance
murD	✓	Experimental	✓	Unknown	1	✓	Peptidoglycan biosynthesis, D-Glutamine and D-glutamate metabolism
murG	✓	Experimental	✓	Cytoplasmic Membrane		✓	Peptidoglycan biosynthesis, Vancomycin resistance
murC	✓	Experimental	✓	Cytoplasmic		✓	Peptidoglycan biosynthesis, D-Glutamine and D-glutamate metabolism
yadF			✓	Cytoplasmic			Nitrogen metabolism
cdsA			✓	Cytoplasmic Membrane	8	✓	Glycerophospholipid metabolism
fabZ	✓	Experimental	✓	Cytoplasmic		✓	Fatty acid metabolism, Biotin metabolism
accA	✓	Approved, Experimental, Nutraceutical	✓	Cytoplasmic		✓	Fatty acid metabolism, Carbon metabolism, Biosynthesis of antibiotics, Microbial metabolism in diverse environments
ribD			✓	Cytoplasmic		✓	Riboflavin metabolism, Biosynthesis of secondary metabolites
ribH	✓	Experimental	✓	Cytoplasmic		✓	Riboflavin metabolism, Biosynthesis of secondary metabolites
fabA	✓	Experimental	✓	Cytoplasmic		✓	Fatty acid metabolism
fabD	✓	Experimental	✓	Cytoplasmic		✓	Fatty acid metabolism
pabC	✓	Experimental	✓	Cytoplasmic		✓	Folate biosynthesis
tmk	✓	Experimental	✓	Cytoplasmic		✓	Pyrimidine metabolism
aroD	✓	Experimental	✓	Cytoplasmic		✓	Phenylalanine, tyrosine and tryptophan biosynthesis, Biosynthesis of secondary metabolites, Biosynthesis of antibiotics
ribE	✓	Approved, Nutraceutical	✓	Cytoplasmic			Riboflavin metabolism, Biosynthesis of secondary metabolites
fabI		Approved, Experimental, Nutraceutical	✓	Cytoplasmic Membrane		✓	Fatty acid biosynthesis, Fatty acid metabolism, Biotin metabolism
ribA			✓	Cytoplasmic		✓	Riboflavin metabolism, Biosynthesis of secondary metabolites
galU	✓	Experimental, Nutraceutical	✓	Cytoplasmic			Amino sugar and nucleotide sugar metabolism, Biosynthesis of antibiotics
pabB	✓	Experimental	✓	Cytoplasmic		✓	Folate biosynthesis
pgsA	✓	Experimental	✓	Cytoplasmic Membrane	4	✓	Glycerophospholipid metabolism
rfbV			✓	Cytoplasmic			
rfbJ	✓	Experimental	✓	Cytoplasmic		✓	Amino sugar and nucleotide sugar metabolism
rfbH	✓	Experimental	✓	Cytoplasmic			Amino sugar and nucleotide sugar metabolism
rfbG	✓	Experimental	✓	Cytoplasmic		✓	Amino sugar and nucleotide sugar metabolism

(continued)

PPI networks are important from perspectives of system level understanding of the cellular processes. Our protein interaction analysis results showed that selected drug targets were involved in crucial cross talks between themselves and there were direct interactions between these proteins. For the selected drug targets, their structures were predicted using RaptorX, where no information was available on crystal structure or homology modeled structure. L-rhamnose, a 6-deoxyhexose is building block of bacterial cell wall, absent from mammals. The rfb gene cluster of *S. typhimurium* LT2 encoding rfbC (dTDP-4-dehydrorhamnose 3,5-epimerase), rfbD (dTDP-4-dehydrorhamnose reductase), rfbF (Glucose-1-phosphate cytidylyltransferase), rfbG (CDP-glucose 4,6-dehydratase), rfbH (Lipopolysaccharide biosynthesis protein RfbH) is essential for expression of O-antigen. Recently peptidoglycan biosynthesis pathway has been extensively subjected to drug targeting. As the enzymes involved in peptidoglycan biosynthesis are indispensable to bacterial survival and do not have any orthologs in humans. glmM, glum, murA to murF, murG and mraY are involved in different stages of peptidoglycan synthesis and indispensable to bacterial cell survival. Folate biosynthesis pathway has significant untapped potential to be exploited as drug targets as the mutations in the enzymes involved in this pathway are reported to produce non-viable phenotypes. folA, folC, folK, folP, pabA, pabB, and pabC are important enzymes of folate biosynthesis pathway. Although fatty acid synthesis pathways are virtually similar between mammals and bacteria, but different profile and arrangements of active sites in these enzymes makes it profoundly trailed pathway for drug targets. fabA and fabZ are essential for unsaturated fatty acid synthesis, but fabZ is considered as the most attractive drug target as it is uniformly expressed in bacterial system while fabA distribution is limited. Riboflavin is a central player in cell physiology as it serves as precursor to many redox reactions as a component of flavin adenine dinucleotide (FAD) and Flavin mononucleotide (FMN). ribD, ribH, ribE were predicted to be good drug targets supplemented with the fact that riboflavin is synthesized *de novo* in bacteria, a pathway absent from mammals.

Conclusion

Drug target and vaccine candidate identification using traditional molecular immunology techniques is a time-consuming, arduous and expensive task. With a myriad of omics techniques, availability of genomic sequence data and novel bioinformatics approaches, our understanding of complex human diseases has radically improved. These techniques have accelerated the pace of drug and vaccine development as they hold the potential for rapid screening and identification of probable vaccine candidates from the whole genome. Our findings may help to develop drug targets and vaccines against diarrheal pathogens once validated experimentally using model organisms.

References

1. Liu L, Johnson Hl, Cousens S *et al*. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *The Lancet* 379(9832), 2151-2161 (2012).
2. Bhutta Za, Das Jk. Global Burden of Childhood Diarrhea and Pneumonia: What Can and Should Be Done? *Pediatrics* 131(4), 634-636 (2013).
3. GBD. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 385(9963), 117-171 (2015).
4. WHO. Diarrhoeal disease. (2013).
5. Collaborators TMDS. Causes of neonatal and child mortality in India: a nationally representative mortality survey. *The Lancet* 376(9755), 1853-1860 (2010).
6. Bushen Oy, Guerrant Rl. Acute infectious diarrhea. Approach and Management in the emergency department. . *Topics in Emergency Medicine* 25(2), 139-149 (2003).
7. *Chapter 19 Diarrheal Diseases*. Keusch G, Fontaine O, Bhargava A *et al* (Ed.^(Eds). Washington (2006).
8. WGO. Acute Diarrhea in Adults and Children A Global Perspective. *J Clin Gastroenterol* 47(1), (2013).
9. CDC. Antibiotic Resistance Threats in the United States, 2013. (2013).
10. Ganguly Nk. Antimicrobial Resistance in Developing Countries
11. Gelband H, Miller-Petrie M, Pant S *et al*. The State of the World's Antibiotics, 2015. (2015).
12. Morgan Dj, Okeke In, Laxminarayan R, Perencevich En, Weisenberg S. Non-prescription antimicrobial use worldwide: a systematic review. *The Lancet Infectious Diseases* 11(9), 692-701 (2011).
13. Woodford N, Ellington Mj. The emergence of antibiotic resistance by mutation. *Clinical Microbiology and Infection* 13(1), 5-18 (2007).
14. Kumar S, Banks Tw, Cloutier S. SNP Discovery through Next-Generation Sequencing and Its Applications. *International Journal of Plant Genomics* 2012 15 (2012).
15. Bharti R, Das R, Sharma P, Katoch K, Bhattacharya A. MTCID: a database of genetic polymorphisms in clinical isolates of Mycobacterium tuberculosis. *Tuberculosis (Edinburgh, Scotland)* 92(2), 166-172 (2012).
16. Liu B, Pop M. ARDB--Antibiotic Resistance Genes Database. *Nucleic acids research* 37(Database issue), D443-447 (2009).
17. Mcarthur Ag, Waglechner N, Nizam F *et al*. The comprehensive antibiotic resistance database. *Antimicrobial agents and chemotherapy* 57(7), 3348-3357 (2013).
18. Vishnoi A, Srivastava A, Roy R, Bhattacharya A. MGDD: Mycobacterium tuberculosis genome divergence database. *BMC genomics* 9 373-373 (2008).
19. Berman Hm, Westbrook J, Feng Z *et al*. The Protein Data Bank. *Nucleic Acids Research* 28(1), 235-242 (2000).

20. Hamosh A, Scott Af, Amberger Js, Bocchini Ca, Mckusick Va. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Research* 33(suppl 1), D514-D517 (2005).
21. Landrum Mj, Lee Jm, Riley Gr *et al*. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Research* (2014).
22. Sherry St, Ward Mh, Kholodov M *et al*. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research* 29(1), 308-311 (2001).
23. Palzkill T. Mitigating Antibiotic Resistance with DNA Sequence Information. Presented at: *Emerging and Persistent Infectious Diseases: Focus on Antimicrobial Drug Resistance*. Baylor College of Medicine, Houston, Texas 2013.
24. Bennett De, Camacho Rj, Otelea D *et al*. Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update. *PLoS ONE* 4(3), e4724 (2009).
25. Harris Ja, Roy K, Woo-Rasberry V *et al*. Directed Evaluation of Enterotoxigenic Escherichia coli Autotransporter Proteins as Putative Vaccine Candidates. *PLoS Neglected Tropical Diseases* 5(12), e1428 (2011).
26. Bagheri S, Mousavi Gargari Sl, Rasooli I, Nazarian S, Alerasol M. A CssA, CssB and LTB chimeric protein induces protection against Enterotoxigenic Escherichia coli. *The Brazilian Journal of Infectious Diseases* 18(3), 308-314 (2014).
27. Butt Am, Nasrullah I, Tahir S, Tong Y. Comparative Genomics Analysis of *Mycobacterium ulcerans* for the Identification of Putative Essential Genes and Therapeutic Candidates. *PLoS ONE* 7(8), e43080 (2012).
28. Ansari Hr, Raghava Gps. Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome Research* 6 6-6 (2010).
29. Grant Ma. Protein structure prediction in structure-based ligand design and virtual screening. *Combinatorial chemistry & high throughput screening* 12(10), 940-960 (2009).
30. Li X, Apel D, Gaynor Ec, Tanner Me. 5'-methylthioadenosine nucleosidase is implicated in playing a key role in a modified futalosine pathway for menaquinone biosynthesis in Campylobacter jejuni. *The Journal of biological chemistry* 286(22), 19392-19398 (2011).
31. Prosser Ga, De Carvalho Lps. Kinetic mechanism and inhibition of Mycobacterium tuberculosis D-alanine:D-alanine ligase by the antibiotic D-cycloserine. *The FEBS journal* 280(4), 1150-1166 (2013).
32. Garmory Hs, Titball Rw. ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. *Infection and immunity* 72(12), 6757-6763 (2004).
33. Harris Ja, Roy K, Woo-Rasberry V *et al*. Directed evaluation of enterotoxigenic Escherichia coli autotransporter proteins as putative vaccine candidates. *PLoS neglected tropical diseases* 5(12), e1428-e1428 (2011).
34. El Zoeiby A, Sanschagrin Fo, Levesque Rc. Structure and function of the Mur enzymes: development of novel inhibitors. *Molecular microbiology* 47(1), 1-12 (2003).

PUBLICATIONS

- **Mehla, K.**, Ramana J. DBDiaSNP: An Open-Source Knowledgebase of Genetic Polymorphisms and Resistance Genes in Diarrheal Pathogens. *OMICS: A Journal of Integrative Biology* 2015; 19(6):354-60.
- **Mehla, K.**, Ramana J. Novel Drug Targets for Foodborne Pathogen Campylobacter jejuni: An Integrated Subtractive Genomics and Comparative Metabolic Pathway Study. *OMICS: A Journal of Integrative Biology* 2015; 19(7):393-406.
- **Mehla, K.**, Ramana J. Structural signature of Ser83Leu and Asp87Asn mutations in DNA gyrase from Enterotoxigenic Escherichia coli and impact on quinolone resistance. *Gene* 2016; 576(1):28–35.
- **Mehla, K.**, Ramana J. Identification of epitope-based peptide vaccine candidates against enterotoxigenic Escherichia coli: a comparative genomics and immunoinformatics approach. *Mol Biosyst.* 2016; 12(3):890-901
- **Mehla, K.**, Ramana J. Molecular Dynamics Simulations and Residue Interaction Network Analysis of Diarrhea associated ETEC gyrA mutants: Repercussions on Quinolone Affinity. [Under Review]
- **Mehla, K.**, Ramana J. Surface Proteome Mining for Identification of Potential Vaccine Candidates against *Campylobacter jejuni*: An In silico approach. [Under Review]
- **Mehla, K.**, Ramana J. Molecular Dynamics Simulations of quinolone resistance associated T86I and P104S mutations in *C. jejuni* gyrA: Unraveling structural repercussions. [Under Review]
- **Mehla, K.**, Ramana J. Tapping into *Salmonella typhimurium* LT2 genome in a quest to explore its therapeutic arsenal: a metabolic network modeling approach. [Under Review]

CONFERENCES

- **Mehla K.**, Ramana J. A Docking Study of Relationship between Point Mutations and Ciprofloxacin Resistance in *Campylobacter jejuni.* **Oral Presentation** at IC LIFE 2014: International Conference on Life Sciences, Informatics, Food and Environment, August 29-30, 2014 at JIIT, Noida
- **Mehla K.**, Ramana J. Constraint based analysis of *Clostridium difficile* metabolic network: FBA and MOMA approach for drug target identification. **Oral Presentation** at International Conference on Advances in Biomedical Engineering, Cancer Biology, Stem Cells, Bioinformatics and Applied Biotechnology (ABECBAB-2016), May 21, 2016 at JNU, Delhi.